Fast Learning in Organizations

Blume, A.

Publication date:
1997

entER
for
omic Research

Discussion paper

[redacted]

R40

organization

games

Tilburg University

Center
for
Economic Research

No. 9765

FAST LEARNING IN ORGANIZATIONS

By Andreas Blume

July 1997

ISSN 0924-7815

Fast Learning in Organizations

Andreas Blume
CentER for Economic Research
Tilburg University

June, 1997

Abstract

This paper explores how efficiency structures language. It starts from the premise that one of language's central characteristics is to provide a means for saying novel things about novel circumstances, its creativity. As such it is a metaphor for the choice of organizational forms that can cope with a changing environment. It is shown how creative language use is achieved via reliance on common knowledge structures, even if those structures are consistent with an *a priori* absence of a common language. *Journal of Economic Literature* Classification Number: C72.

KEYWORDS: Language, Coordination, Optimal Learning, Common Knowledge, Group Theory

---

* This paper was begun while I enjoyed the hospitality of the MEDS department at Northwestern University. The first version was completed at the CentER for Economic Research at Tilburg University. I am grateful to David Austen-Smith, Eddie Dekel, Oguz Durumeric, Ehud Kalai, Bart Lipman, Alejandro Manelli and Roger Myerson for helpful discussions and to Vince Crawford, Hans Haller and Joel Sobel for comments. I have benefitted from comments by seminar participants at the CentER for Economic Research (Tilburg University), CREED (University of Amsterdam), University of Bielefeld, Humboldt University, University of Saarbrücken, and the SAET conference in Antalya.

Drug firms need to get their inventions on to the market quickly. That is easier when researchers and factory designers talk to each other. ... As competition grows even more fierce, more companies may try to make sure that their boffins and factory designers talk to each other early on. For the money it saves, it might even be worth paying for an interpreter. [The Economist, November 9th 1996]

1 Introduction

This paper explores how efficiency structures language. It starts from the premise that one of language's central characteristics is to provide a means for saying novel things about novel circumstances, its creativity (e.g. Aitchison [1993]). As such it is a metaphor for the choice of organizational forms that can cope with a changing environment.

We are looking for structure in the mode of acquisition of a lexicon. The structure arises in the form of an *optimal learning rule* in a coordination game that is played over time. Initially the players' ability to coordinate is limited by an absence of a common language. A language is developed over the course of the game via observations of past play. The players' attainable strategies are constrained by how far their language has evolved at each point. How fast the language evolves depends on which equilibrium in attainable strategies is played. The primary aim of the paper is to show that *a priori* absence of a common language does not preclude rapid learning and creative use of language.

Two interpretations for the origins of efficient structures are available, an evolutionary or learning interpretation and a mechanism design interpretation.[1] Under the former there is evolutionary pressure on behaviors that depend on factors that are constant across novel circumstances. Since in each novel situation, the novelty is dealt with by learning, a complete model along these lines would be one of learning how to learn. The mechanism design interpretation on the other hand asks directly for learning rules that cope efficiently with novelty.

In our model the origin of fast learning and creative use of language is structure in the prior beliefs. This structure itself does not constitute a common language, as words in the lexicon need not have an *a priori* meaning, but may facilitate the acquisition of a lexicon once certain parameters are determined via observations made by the players over the course of the game.

We will capture novelty of a set of objects, tasks, types, messages etc.

[1] Rubinstein [1996] calls the designer of a such a mechanism a "linguistic engineer."

via an *absence of a common description of that set.* On occasion only some aspects of a set need to be novel. Then those aspects will lack a common description while others may have a common description.

Let Ω be a finite set and let us try to be explicit about players' *descriptions* of that set and about what players know about each others' descriptions. We will follow Crawford and Haller's [1990] (CH in the sequel) suggestion that descriptions of Ω be permutations of the set Ω itself, and that in case players are ignorant about another player's description they have a uniform belief over alternative permutations. In the case where Ω is part of the description of a game, this approach permits us to model the absence of a common description in a standard Bayesian game setting.

Part of our intention in this paper is to point out that absence of a common description of certain elements in a game is compatible with a rich set of optimal learning rules. The key to our results is a careful reexamination of CH's notion of absence of a common language. They effectively assumed that *all* permutations of the set Ω are equally likely.[2] In that case a common language can be learned only very slowly, observation by observation. However, there are a great many coordination problems in which there is no scope for coordination on a priori grounds, and yet learning can take place very rapidly. Also there are a great many coordination problems in which it is imperative that a rich language be learned as opposed to just a single action being identified.

To us, following CH's approach, absence of a common description means that no subset of Ω is distinguished by the common-knowledge distribution ϕ over the set of descriptions that expresses the players' uncertainty. If all descriptions are possible and all descriptions are equally likely, then we have indeed absence of a common description. Equal probability we clearly cannot dispense with. However, it is less obvious that we need to invoke *all* possible descriptions of a set. A five-element set can be ranked in 120

[2] This is the assumption used in the body of CH's paper. In the appendix they consider a more general formulation that is compatible with the approach in the present paper. There the extent of a common description of Ω is expressed via a partition of Ω, with absence of a common description captured by the trivial partition. The partition is refined as a function of the history of play, where a very general form of history dependence is possible. CH do not closely investigate the relation between histories and partitions. For example, for two elements of Ω not to be commonly distinguished it must be the case that the partitions they induce, if we exchange them in otherwise identical histories, must be identical up to symmetry. The approach of the present paper ensures that this is the case, in a setting where beliefs are expressed as permutations of the analyst's ranking of the elements of Ω.

different ways. We want to suggest in this paper that absence of a common description can be expressed by nontrivial subsets of the set of 120 descriptions. In the next section we will give examples to suggest that often such subsets are more natural than the entire set of possible permutations.

Clearly, not every subset of the set of permutations will be a candidate for expressing absence of a common language. A singleton set for example expresses the exact opposite, a language in which each element can be distinguished. Less trivially, the candidate set may have properties that allow the players to use introspection to further reduce the set. The first major result of the paper (Proposition 1) is a characterization of sets of permutations that do not allow inferences that would further reduce the set. Irreducibility of a set of permutations turns out to be equivalent to requiring the set to form a *permutation group,* a subgroup of the *symmetric group* of all permutations of Ω, denoted by S_Ω. This is quite natural given that irreducibility expresses symmetry among the permutations in the set; no subset can be singled out. Groups are the way to express symmetry in mathematics, e.g. the symmetries of a geometric figure.

Irreducibility of the set of descriptions in the support of the players' beliefs is necessary but not sufficient for "absence of a common description." Again, a singleton is irreducible; if we make it the standard description, a singleton becomes the identity, a trivial group. Absence of a common description requires not only that one cannot make distinctions among the descriptions in the support of the players beliefs but also that this set of descriptions does not distinguish among the elements of Ω itself. Such distinctions arise naturally because every group that acts on Ω induces a partition of Ω. Only if our group of descriptions induces a trivial partition, composed only of Ω itself, do we say that the set of descriptions satisfies "absence of a common description."

If the set Ω is part of the description of a game, then it is likely that the players over the course of the game make observations about Ω. We identify such observations as subsets of Ω and note that they introduce distinctions among the elements of Ω. At the very least the observed subset becomes distinct from its complement but in general (this is the point of this paper) there will be further distinctions. These additional distinctions arise because an observation singles out not only a subset of Ω but also a subset of the set of descriptions of Ω; roughly, it focusses on all those descriptions that describe the observation in a particular way (e.g. if the observation is a singleton, it may pick all those descriptions that assign rank one to the observation).

One question that arises in this context is whether the set of descriptions that are selected by an observation is again irreducible. That this is indeed the case is shown in Proposition 2.

The set of all descriptions of a set Ω satisfies irreducibility and absence of a common description. Learning in this environment is slow because at each step only the observation itself becomes identified. The question arises then whether there are other sets of description that satisfy irreducibility and absence of a common description and can be learned faster. Proposition 3 takes a first look at this question. It deals with observations that are pairs of elements of Ω. This would be appropriate for example in a two player coordination game in which the positions of the players are not distinguished. Proposition 3 shows that for any finite set Ω there exists a group characterized by absence of a common description such that for large Ω almost every observation of a pair of elements of Ω induces a common description of all of Ω. If we apply this to repeated coordination games, then two differences with CH emerge in this case: (1) Typically coordination will be achieved in the second round, and (2) whenever coordination is achieved, then at the same time all other actions become identified. The latter is inconsequential in repeated coordination games but can easily become essential once we look at other classes of coordination problems.

Proposition 4 shows that in some sense the fast learning phenomenon identified in Proposition 3 is ubiquitous. Every subgroup of S_Ω can be learned faster than the symmetric group.

The remaining two results examine the nature and variety of fast learning. Proposition 5 shows that there are definite restrictions on the form that optimal learning can take if we represent players' uncertainty through permutations. In general one can rule out learning paths in which for a while only the observations themselves become identified and then suddenly all elements of Ω become identified. Therefore, in general, fast learning will involve observations that identify more than the observation itself but less than all of the elements of Ω.

The final result, Proposition 6, shows that for any lower bound on the size of Ω, there is a rich set of groups representing absence of a common description that can be learned with any prespecified number of observations.

2 Strategies, Games and Examples

This section recalls and generalizes the definition of an *attainable strategy*. It then discusses a series of different coordination games to motivate optimal learning and to demonstrate the role of fast learning.

If we take Ω as part of the description of a game, it might represent actions of players, some part of their private information, or even the players themselves. Strategies, as usual, are functions that map the players' information into their actions. Thus Ω might be either part of the domain or of the range of these functions; there might even be multiple sets Ω_j representing different aspects of the environment for which the players lack a common language description. The fact that players lack a common language description of Ω, formally appears as a restriction on their beliefs. Given his own description, a player's strategy is unrestricted. However, since a common knowledge description is not available, the other players' beliefs assign equal probability to all strategies of a player that differ only in the treatment of elements of Ω that are not commonly distinguished.

This approach permits us to model the absence of a common description in a standard Bayesian game setting. Players' descriptions are drawn from a common-knowledge distribution ϕ on the set of possible descriptions of Ω, where we take this set as the set of permutations of the elements of Ω. The description drawn by a player becomes that players' private information. Absence of a common language description means that players' strategies depend on their private information only in a restricted sense. Say, a player has a certain strategy given one description; then his strategy in terms of another description is exactly the same, conditional on the description (e.g. the j's action in one description plays exactly the same role as the j's action in another description).

A player's strategy may well depend in great detail on the set Ω. However from the perspective of the other players it can only depend on those aspects of Ω that have a common description. If therefore we adopt the convention that a player's strategy expresses the beliefs other players hold about her (see Rubinstein [1991]) then lack of common describability can be expressed in terms of restrictions on players' strategies. Therefore we call a strategy *attainable* for a player if it satisfies the condition that he randomizes uniformly over all pure strategies that differ only in the treatment of elements of Ω that are not commonly distinguished.

Among the attainable strategy profiles we are interested in those that maximize the players' ex ante payoff. These we call *optimal attainable strate-*

gies. Note that we deliberately ignore higher order coordination problems that may arise if there are multiple optimal attainable strategies. In the spirit of the introduction we think of those multiplicities as being eliminated either by evolutionary pressures or by Rubinstein's [1996] "linguistic engineer."

Next we discuss three motivating examples.

2.1 Repeated Coordination Games

Here we recapitulate a central insight from the work of Crawford and Haller. Consider the following game played on a finite set Ω: Each of two players chooses one element of Ω simultaneously and independently. If both make identical choices, then their payoffs equal 1, otherwise their payoffs are equal to 0. Let this game be repeated infinitely often, and let repeated game payoffs be equal to the discounted sum of stage game payoffs, with a discount factor $0 < \delta < 1$. Let the players' uncertainty be represented by a uniform distribution over all possible permutations of the analyst's description of Ω. Also, let the players' positions be not distinguished, which forces them to use identical strategies. Then, as shown by CH:

1. If $\#\Omega = 2$, then there is an essentially unique optimal attainable strategy. In the first stage absence of a common description requires that players randomize uniformly over their two actions. In subsequent stages, if players have not yet succeeded to coordinate, they switch their action in each period with probability one half. Once players achieve coordination in some period, they can maintain coordination in subsequent periods by simply continuing to use the same action (here we have inessential nonuniqueness because once actions are identified, there are many ways of sustaining cooperation).

2. With $\#\Omega \geq 6$ the following is an optimal attainable strategy, quoting CH: "Play each of your actions with equal probability in the first stage. If coordination results, maintain it by repeating your first-stage action. If not, rule out all of your actions but your first-stage action and the action that would have yielded coordination, given your partner's first-stage action. Then play the resulting 2×2 game using your part of its optimal attainable combination."

Note that the expected coordination time is $\hat{t} \geq 2$, that coordination sometimes will take a long time, and that even if coordination is achieved the players are in general far from having developed a common language. In that sense the optimal coordination process is slow and of limited scope. For the game at hand the limited scope is of no consequence, however in other contexts learning a common language fully may be essential.

2.2 A Rudimentary Grammar

Consider the following game played repeatedly between two players, a sender and a receiver. At the beginning of the game the sender learns his private information and sends a message to the receiver. Upon receiving the message the receiver takes an action. Payoffs depend only on the sender's private information, his type, and the receiver's action.[3] The payoff to both players is one if the receiver's action matches the sender's type and zero otherwise. There is exactly one matching action for each type of the sender. Messages do not affect payoffs directly. Assume that after each round the players commonly observe the type drawn for that round, the message sent in that round, and the action taken. The sender's private information is determined anew in each round according to a uniform distribution.

Additional structure is provided by types and messages being strings. To simplify the discussion let types be triples formed by permutations of the letters A, B and C, e.g. (B, C, A), and let messages be triples as well, formed by permutations of $*, \#$ and $\&$. To rule out *a priori* focal points we will take this to be the analyst's description and assume that for each player there exists a private set of six symbols, taking the roles of $A, B, C, *, \#$ and $\&$, not known to the other player.[4] Now assume that either through prior experience with a similar environment or handed down by the linguistic engineer the players have a common knowledge ranking of permutations of three objects; i.e., given a particular ordered triple of objects, they can generate a common-knowledge sequence of all six ordered triples of those objects.

The first time the game is played, the sender and the receiver lack a common-knowledge description of the types space and of the message space. According to our definition of attainable strategy, each type randomizes uniformly over all messages and the probability of a matching action being taken is $1/6$. In the following round however each player, in terms of his own representation, has observed two triples and constructed two corresponding

[3]Wärneryd [1993], Blume, Kim and Sobel [1993] and Blume [1996] consider the evolution of meaning of *a priori* meaningless messages in sender receiver games.

[4]The focal point notion was first proposed by Schelling [1960].

common knowledge rankings of those triples. Thus every type and every message has become identified, meaning that there is an attainable strategy guaranteeing successful communication in the second round and thereafter.

Crucial here is access to a common knowledge rule for generating permutations of triples. This in effect puts the elements of the type set like beads on a circle with an orientation; once the initial bead is fixed, the rank of all other beads is determined. The initial uncertainty about the 6-element type (message) set is then represented by a particular cyclic subgroup of the symmetric group S_6. Finally, note that the common knowledge rule can be interpreted as implementing a very simple and natural (common knowledge) rule that uses the first observation to generate a bijection between the sets $\{A, B, C\}$ and $\{*, \#, \&\}$.

Contrary to the first example, coordination is guaranteed from the second period on and a complete common language is learned. Moreover players use their language creatively in that in the second period they are likely to indicate a novel type (not observed before) via a novel message (not sent before). Batali [1996] has referred to similar structures as *grammars*, emphasized the role of such grammars for the expression of novel meanings and inquired into the evolution of such grammars.

2.3 Coordination on Spheres

The following example combines features of the first two and also indicates how the linguistic engineer could be dispensed with. Let Ω be the set of points on a sphere (with no further distinctions among those points). In each period two agents are chosen to play a repeated game on the sphere. Each period is divided into rounds. In each round, the two chosen agents simultaneously and independently pick a location on the sphere. Once both players have made their choices, the choices are revealed to the agents. If the locations picked are the same, then their payoff in that round is one, otherwise, it is zero. In each period the two chosen agents receive the discounted sum of their payoffs from each round. In the following period a new pair of agents is drawn from the population and the same game is played with those two agents on a new sphere. Each pair of agents can observe all the choices made by their predecessors on their respective spheres.

The game played within a period is quite similar to the game played in our first example. It is a repeated coordination game. The probability of coordination in the first round equals zero, the same as the limit of first-round coordination probabilities in the first example if we let the number of locations go to infinity. However, there exists an optimal attainable strategy that guarantees coordination in the second round (and thereafter) with probability one. One such strategy prescribes that in the second round the players choose the midpoint of the shortest distance between their first-round choices. This midpoint is almost always well defined, the exception being antipodal first-period location choices.

Unlike in the first example we can guarantee coordination in the second round. As in the second example this is due to the common-knowledge structure, here of the space of locations. Once we speak of a sphere, we implicitly limit the set of permutations of locations to be considered to motions, those are all permutations of points on the sphere that leave distances invariant.

The repetition of the repeated coordination game serves only to illustrate how one may think of dispensing with the linguistic engineer. If we think of agents in different periods as experimenting with different strategies, and adopting strategies that were successful for their predecessors, we may well expect convergence to an optimal attainable strategy over time.

3 Irreducibility of Descriptions

In this section we formulate a desideratum for irreducibility of a set of descriptions of a finite set Ω and show that a set of descriptions satisfying this condition must be a subgroup of the symmetric group S_Ω. Let $\#\Omega = n$ and let H be a set of rankings of the elements of Ω; i.e., each $h \in H$ is a bijection $h : \Omega \to \{1, 2, ..., n\}$. It is without loss of generality to identify the set Ω with $\{1, 2, ..., n\}$ (and we will do so freely in the sequel), in which case each $h : \Omega \to \Omega$ becomes a permutation of Ω. Furthermore, it is without loss of generality to choose the analyst's labelings such that they correspond to one of the rankings; i.e., $\exists e \in H$ such that $e(\omega) = \omega, \forall \omega \in \Omega$. That is, from the analyst's perspective we can make one of the rankings the standard. Concerning permutations, the standard is then expressed in terms of itself simply via the identity function $e(\cdot)$. All this means is that we can choose one of the rankings considered possible by the agents as the standard ranking.

However, it is very important to emphasize that we do not want the agents to be able to single out particular rankings by introspection. This requires that the agents do not know the analyst's ranking. But it also requires that rankings are essentially indistinguishable. No matter which standard we choose, the set of rankings H expressed in terms of that standard must always be the same.

Suppose for example that we replace the current standard ranking by an alternative ranking $g \in H$. In that case an object that was originally labeled ω is now labeled $\omega' = g(\omega)$. We can then express all other rankings, f, in terms of the new standard, i.e.

$$f(\omega) = f \circ g^{-1}(\omega').$$

Thus $f \circ g^{-1}$ (or abbreviated fg^{-1}) becomes the new representation of what used to be the ranking f.[5]

Our condition that the set of rankings H not change if we choose a particular one as the standard then amounts to requiring that

$$\{fg^{-1} | f \in H\} = H \quad \forall g \in H,$$

or in abbreviated form

$$(IA) \qquad Hg^{-1} = H \quad \forall g \in H.$$

We will refer to this condition as the irreducibility axiom (IA). If the players' uncertainty over the description used by the other players is expressed in terms of H, then the irreducibility axiom guarantees that the players cannot introspectively reduce their uncertainty by adopting one particular standard from the set of all possible standards.

The irreducibility axiom imposes a strong condition on sets of possible descriptions. The implications of imposing IA are explored in the following simple result. This proposition characterizes irreducible sets of descriptions as *permutation groups*. For that purpose recall that a group G consists of a set (here the set of permutations), an operation ,"$\circ$", on the set (here composition of permutations) such that for all $f, g, h \in G$, (i) the associative law holds ($f \circ (g \circ h) = (f \circ g) \circ h$); (ii) there exists an identity element e such that $e \circ g = g = g \circ e$; and (iii) for all $g \in G$ there exists an inverse g^{-1} with $g \circ g^{-1} = e = g^{-1} \circ g$.

One group that is of particular interest to us is the set of *all* permutations of elements of the set Ω, denoted by S_Ω. This group is referred to as the symmetric group on Ω. If G is a group and H is a subset of G, then H is

[5] Note that we adopt the convention that permutations "act on the left," i.e. whenever we compose permutations, the rightmost in the product is performed first, followed by its neighbor on the left and so on.

a subgroup of G, denoted $H \leq G$, if H contains the identity, and is closed under taking inverses and under the group operation.

Proposition 1 *A set of descriptions H of a set Ω satisfies the irreducibility axiom (IA) if and only if H is a subgroup of the symmetric group S_Ω.*

Proof: Consider necessity first. (i) (existence of the identity). Earlier we found it convenient to simply assume that $e \in H$. However, this is also implied by (IA) because $Hg^{-1} = H$ and $g \in H$ together imply that $e = gg^{-1} \in H$. (ii) (existence of an inverse) $Hg^{-1} = H$ and $e \in H$ together imply that $g^{-1} = eg^{-1} \in H$. (iii) (composition of permutations in H defines an operation on H) ($Hg^{-1} = H, \forall g \in H$) and ($g^{-1} \in H, \forall g \in H$) implies that $Hg = H, \forall g \in H$ from which it follows that $fg \in H, \forall f, g \in H$. (iv) (associative law) associativity is inherited from S_Ω. Conversely, if H is a permutation group, then since H is closed under the group operation, we have $Hg \subset H, \forall g \in H$. Since $e \in H$, this implies that $Hg = H, \forall g \in H$, and in conjunction with $g^{-1} \in H, \forall g \in H$, we also have $Hg^{-1} = H, \forall g \in H$. □

4 A Priori Distinctions

The agents may *a priori* be able to commonly distinguish some subset $\Delta \subset \Omega$ from other subsets of Ω. If we express the agents' beliefs via a set of possible descriptions, H, then H must reflect the agents' ability to make such common distinctions. This is indeed the case because H associates a set of possible labels with each element $\omega \in \Omega$. Points ω for which these sets of labels differ from each other can be commonly distinguished by the agents. To make all this precise, it is useful to introduce the notion of a group G acting on a nonempty set Ω. For each $\omega \in \Omega$ and for each $g \in G$, define an element $g(\omega) \in \Omega$. Then the group G *acts on* Ω if

(i) $e(\omega) = \omega \;\; \forall \omega \in \Omega$, and

(ii) $h(g(\omega)) = hg(\omega) \;\; \forall \omega \in \Omega$, and $\forall\, g, h \in G$.

In our case, we are only interested in permutation groups, $G \leq S_\Omega$ and their "natural" group action on Ω, where each permutation is viewed as a function. Given any group G acting on Ω, we can associate with any $\omega \in \Omega$ the image of ω under the group action. Define the *orbit* of ω under G acting on Ω as

$$\mathcal{O}(\omega) := \{g(\omega) | g \in G\} \subset \Omega.$$

Then points ω and ω' can be distinguished if and only if they belong to different orbits. In that regard, it is useful to know that the set of orbits forms a partition of Ω (e.g. Rotman [1996], p.122).

5 Absence of a Common Description

We can consider the special case of a single orbit. Then $\mathcal{O}(\omega) = \Omega \ \ \forall \omega \in \Omega$. In this instance none of the elements of Ω can be distinguished from any other element of Ω. We say then that there is an "absence of a common description." A group H with only a single orbit is called *transitive* (e.g. Dixon and Mortimer [1996], p.8).

6 Observations

For the moment assume that the initial beliefs are characterized by an absence of a common description of Ω, i.e. by a transitive permutation group acting on Ω. We want to examine how the agents can update and coordinate their beliefs on the basis of observations about Ω. The key here is that the agents can use their observations to label points in Ω.

Different labeling rules could arise or be suggested by the linguistic engineer ("name the first object observed 5," "name it 7"). Note however that because these rules are distinguishable, agents can arrive at a common rule by introspection, or if necessary, through repeated encounters of similar situations. From the analyst's perspective, it is then without loss of generality to consider a particular rule. We will adopt the rule that the first element observed is named "1." And as long as observations do not restrict the manner in which the remaining elements are named, the k-th element observed is named "k."

Thus far we have covered the case of singletons observed in sequence. We can generalize this idea to observations of sets of elements; observation of a set amounts to simultaneous observation of the points in the set in a manner that does not permit distinctions among points in the set. Suppose the set $\Delta \subset \Omega$, with $\#\Delta = m$, is observed initially.[6] For any such m, $1 \leq m \leq n$, consider sets of the form $\{l_1, l_2, \ldots, l_m\}$ with $l_i \in \{1, \ldots, n\}$, and $l_i \neq l_j$ for

[6] If Δ is an observation, then one may worry whether the agents might not be better off if they took the set $\Omega \setminus \Delta$ as the basis for their introspection. We will later see that this is not an issue.

$i \neq j$. For any fixed m and n consider some arbitrary complete order, $\succ$, on these sets (such that any collection of such sets has a maximal element).

Given the observation of a set $\Delta \subset \Omega$ (expressed in terms of some standard $h \in H$), we can find for each $g \in H$ the set

$$\Delta^g := \{g(\omega) \in \Omega | \omega \in \Omega\}$$

and construct the collection of sets

$$C(\Delta) = \{\Delta^g | g \in H\}.$$

From above, $C(\Delta)$ has a $\succ$-maximal element, say Δ^*; note that in general there will be multiple g such that $\Delta^g = \Delta^*$. Pick some $g^* \in H$ such that $\Delta^* = \Delta^{g^*}$ and reexpress the elements of Ω in terms of g^*, i.e. $\omega' = g^*(\omega)$. In terms of these new variables we then have functions $g' = gg^{*-1}$ and $H' = Hg^{*-1}$.

Define the set of permutations

$$H'_{\{\Delta^*\}} = \{g' \in H' | \Delta^{*g'} = \Delta^*\}.$$

In group theory, this is referred to as the *setwise stabilizer* of the set Δ^* (e.g. Dixon and Mortimer [1996], p.13). Intuitively, this is the set of all descriptions that do no alter the set of labels for elements of Δ^*, where Δ^* is the observation Δ expressed in terms of the new coordinates. To summarize, we used the order $\succ$ to pick a set of labels for the observation Δ, identified all permutations which assign the same labels to Δ and used one of them to define new coordinates. In terms of these new coordinates, the permutations selected by the observation are the setwise stabilizer of the observation. At this point, while using the new coordinates, it is convenient to simply drop all the $*$'s and primes and to consider

$$H_{\{\Delta\}} = \{g \in H | \Delta^g = \Delta\}.$$

One checks easily that for any Δ, the setwise stabilizer $H_{\{\Delta\}}$ is a subgroup of H. Thus, we have the following result.

Proposition 2 *Given an irreducible set of description $H \leq S_\Omega$, an observation $\Delta \subset \Omega$ induces an irreducible set of descriptions $H_{\{\Delta\}}$.*

We derived this result under the assumption that H induces a single orbit. However, it is valid more generally. If Δ is a subset of one of the orbits of H, the preceding argument applies unchanged. If instead Δ intersects multiple orbits, then each of these intersections can be treated as a separate observation.

At this point it is useful to return to the question of whether agents might prefer to update the support of their beliefs based on $\Omega \setminus \Delta$ rather than Δ. For that purpose note that when a group G acts on Ω, then each function $g(\cdot)$ is a bijection since its inverse is simply $g^{-1}(\cdot)$, where g^{-1} is the inverse of g in the group G. It is easily seen that if each $g(\cdot)$ is a bijection, then

$$H_{\{\Delta\}} = H_{\{\Delta^C\}} \quad \forall \Delta \subset \Omega.$$

7 Learning to Coordinate

The next question we want to address and the central theme of this paper is how hard it typically is to solve two-player coordination problems.

A similar question was posed before by Crawford and Haller [1990]. They were primarily concerned with a particular form of the absence of a common description where the players' initial beliefs take the form of a uniform distribution over all possible permutations.[7] In considering this case, they did not only rule out the possibility for *a priori* coordination, they also, implicitly, assumed that each observation identified merely the observation itself, and did not lead to any further differentiation of the set Ω. In our framework this can be accomplished by letting the permutation group G that expresses the players' initial beliefs simply be equal to the symmetric group S_Ω.

One of the central insights of CH's paper dealt with the case with extreme symmetry, where positions of players were not distinguished which meant that a player's strategy could not depend on the players identity. They showed that with sufficiently many actions, there is an essentially unique optimal way to play a repeated coordination game: In the initial round, each player's action choice is arbitrary. If the players manage to coordinate in the first round, they stick to their initial action choices forever after. Otherwise they randomize uniformly over two actions, their own initial action, and the action that would have been optimal against the choice of the other player. Randomization continues until coordination is achieved. Thereafter the players stick to their coordinating actions.

[7] In an appendix they consider a general formulation using partitions but do not investigate optimal learning in this case.

The optimal coordinating process in the case where the players' uncertainty is expressed as a uniform distribution over all permutations is slow and limited. Each observation identifies merely the observation itself and once coordination is achieved, at most three sets of actions are identified, the coordinating action, the unsuccessful action and the unused actions. In summary, players start out without a common language and end with a partially common language that is just rich enough to make coordination possible. No common distinction among unused messages arises; players do not use their language creatively.

The introductory examples of a rudimentary grammar and coordination on spheres show that absence of a common description need not imply that learning is as slow and limited in scope as in the case where all permutations are considered. Our next result examines the intuition underlying these examples formally, in a finite setting.

To state our result we will need a little more group theoretic terminology. If $g \in S_\Omega$ and $\omega \in \Omega$, then we say that g *fixes* ω if $g(\omega) = \omega$; otherwise g *moves* ω. Let $\omega_1, \ldots, \omega_r \in \Omega$, $\omega_i \neq \omega_j$ for $1 < i, j \leq r$, $i \neq j$. If g fixes ω_j for $j \neq 1, \ldots, r$, and if $g(\omega_1) = \omega_2, g(\omega_2) = \omega_3, \ldots, g(\omega_{r-1}) = \omega_r, g(\omega_r) = \omega_1$, then g is called an *r-cycle*, and is denoted $(\omega_1, \omega_2, \ldots, \omega_r)$.

Let G be a group and $g \in G$. If $g^k = e$ for some $k \geq 1$, then the smallest such k is called the *order* of g. If G is a group and $g \in G$, then $< g >:= \{g^n | n \in \mathcal{Z}\}$ is the *cyclic subgroup* of G that is *generated* by g. One convinces oneself easily that if g is an r-cycle, then g has order r and $< g >= \{e, g^1, \ldots, g^{r-1}\}$.

When positions in a two-player simultaneous-move game are not distinguished, then observations of the two players' simultaneous actions are not distinguished and therefore we must consider the corresponding setwise stabilizers. We will refer to observations of k-element sets as "k-observations." The next result on the effect of two-observations is applicable to two-player games in which players' positions are not distinguished.

Proposition 3 *Let $G =< g >$ where g is an n-cycle and consider the natural action of G on Ω. Then*

1. *G expresses absence of a common description;*
2. *if n is odd, then every two-observation induces a common description of all elements of Ω; and,*
3. *if n is even, then a proportion $\frac{n-1}{n}$ of all possible two-observations induce a common description of all elements of Ω.*

Proof: For (1) to hold, G must have a single orbit. The definition of an n-cycle implies immediately that G does indeed have a single orbit.

For (2) let n be odd and consider, without loss of generality, the two-observation $\{\omega_1, \omega_2\}$. Since g is an n-cycle, e is the unique element $h \in G$ such that $h(\omega_1) = \omega_1$. To derive a contradiction, suppose that $G_{\{\{\omega_1,\omega_2\}\}}$ is not a singleton. Then there must be numbers k, l, $0 \leq k, l \leq n-1$ such that

$$g^k(\omega_1) = g^l(\omega_2),$$

and

$$g^k(\omega_2) = g^l(\omega_1),$$

which can be restated as $g^{k-l}(\omega_1) = \omega_2$ and $g^{k-l}(\omega_2) = \omega_1$. Therefore $g^{2(k-l)}(\omega_2) = \omega_2$. Let $k \geq l$ without loss of generality. Then, since g is an n-cycle by assumption, n must be a divisor of $2(k-l)$ and since $2(k-l) < 2n$, it must be the case that $2(k-l) = n$. This implies that n is even, thus generating a contradiction. Therefore $G_{\{\{\omega_1,\omega_2\}\}}$ must be a singleton.

The argument that we just gave to show that $G_{\{\{\omega_1,\omega_2\}\}}$ must be a singleton whenever n is odd, clearly does not work for the case where n is even, and indeed the claim is not true when n is even. However, it remains true that any two-observation $\{\omega_1, \omega_2\}$ whose stabilizer is not a singleton satisfies

$$g^{\frac{n}{2}}(\omega_1) = \omega_2.$$

For all other two-observations of the form $\{\omega_1, \omega\}$ with $\omega \neq g^{\frac{n}{2}}(\omega_1)$, $G_{\{\{\omega_1,\omega\}\}}$ is a singleton. That is, $n-1$ out of n two-observations $\{\omega_1, \omega\}$ have the property that $G_{\{\{\omega_1,\omega\}\}}$ is a singleton and thus signifies a common description. □

This proposition deals with a class of subgroups that capture a similar intuition as "coordination on spheres." This class could be referred to as "coordination on a directed circle." Note that if Ω represents actions in a coordination game, then here acquisition of a common language is (almost always) assured in the second round and the players achieve a complete common knowledge description of Ω. One difference with the spheres example is that there second-round coordination can be guaranteed but not a complete common language labeling of all points on the sphere in the second round.

When positions are distinguished in a two-player simultaneous-move game, the two players' simultaneous actions become separate singleton observations. If the game in question is a simple coordination game and we are interested in how players achieve coordination through repeated play, then learning with distinguished positions is essentially trivial; since positions are distinguished, we can simply assign the (commonly known) label "player one" to one of the players. There is then an attainable strategy profile in which player one repeats his first-round action in subsequent rounds and player two uses a best reply to player one's action in all rounds following the first round. This is independent of which subgroup of the symmetric group S_Ω expresses the players' uncertainty at the beginning of the repeated game.

Learning with distinguished positions becomes less trivial if we alter the game. For example, consider a game in which, as before, players receive a positive payoff if and only if they meet in some location but once a locations has served as a meeting place it cannot do so again until at least k rounds have passed, where $1 < k \leq n-1$. Also, simplify by letting only player one's action be commonly observable after each round. If the initial uncertainty is described by S_Ω, then locations become identified only by player one's choice of those locations. Successful coordination on some location in the first period, for example, does not guarantee coordination in subsequent periods because that location cannot be revisited for some time and because for the other locations any kind of common description is still lacking.

Especially for large n and k, coordination in the initial phase of the game becomes quite tedious. Consider as an extreme case $k = n-1$ then sustained coordination is possible but in order to achieve it, the players need to acquire a complete common description of the set of locations Ω. Even if the players cared about nothing else but achieving such a common description as early as possible, it would still take $n-1$ periods. Given, that players discount the future and given that coordination is a chance event at any location that is not yet commonly described, the expected time until a full common description is achieved exceeds $n-1$. For example, if agents do not coordinate in the first period, then discounting will induce them to both visit player one's first period choice in the second period. Thus, no new location is identified in the second period.

In this example the full description of Ω is acquired very slowly, one observation at a time. This contrasts with the case where the initial uncertainty is described by $< (\omega_1, \omega_2, \ldots, \omega_n) >$ and where therefore coordination can be guaranteed in all rounds but the first round.

For the remainder of the paper we will concentrate on the case of singleton observations, as in the example with distinguished positions. Obviously, it is always the case that a full description of a set Ω with $\#\Omega = n$ can be acquired with no more than $n-1$ observations. The example shows that fewer observations may suffice. Indeed, the following result shows that the case where $n-1$ observations are needed is in a certain sense atypical. Whenever absence of a common description of an n-element set Ω is captures by a nontrivial subgroup G of S_Ω, denoted $G < S_\Omega$, strictly fewer than $n-1$ observations are needed to establish a common description of Ω.

For the next result, since it is concerned with the accumulation of observations it will be useful to have ready the definition of the *pointwise stabilizer* $G_{(\Delta)}$ of a set Δ (e.g. Dixon and Mortimer [1996], p. 13):

$$G_{(\Delta)} = \{g \in G | \omega^g = \omega, \forall \omega \in \Delta\}.$$

Proposition 4 *Let $\#\Omega = n$ and let the uncertainty over Ω be described by the group $G < S_\Omega$. Then Ω can be learned with strictly less than $n-1$ observations.*

Proof: If G is not transitive on Ω, then we can partition Ω into the orbits of G on Ω (e.g. Rotman [1996], p.122). Let there be $r > 1$ such orbits and denote the $\rho's$ orbit by $\mathcal{O}_\rho$. Then $n = \sum_{\rho=1}^{r} \#(\mathcal{O}_\rho)$. Since each orbit can be learned with at most $\#(\mathcal{O}_\rho) - 1$ observations, Ω can be learned with at most $\sum_{\rho=1}^{r}(\#(\mathcal{O}_\rho) - 1) = n - r$ observations. Essentially the same argument works if there is a set Δ of $l < n-1$ observations such that the pointwise stabilizer $G_{(\Delta)}$ fails to be transitive.

Consider therefore the case where after any number $l \leq n-1$ of observations, given by Δ, the pointwise stabilizer $G_{(\Delta)}$ remains transitive on $\Omega \setminus \Delta$. We know that in general the following relationship between orbits and stabilizers holds (e.g. Rotman [1996], p.123):

$$\#(G_\omega) = \frac{\#(G)}{\#\mathcal{O}(\omega)}, \forall \omega \in \Omega.$$

Thus, if G is transitive on Ω, we have

$$\#(G_\omega) = \frac{\#(G)}{n}, \forall \omega \in \Omega.$$

As long as after each additional observation the stabilizer remains transitive on the complement of the set of observations, it follows by induction that the pointwise stabilizer of the l-element set Δ satisfies the condition

$$\#(G_{(\Delta)}) = \frac{\#(G)}{(n!/(n-l)!)},$$

and if we set $l = n - 1$, we have

$$\#(G_{(\Delta)}) = \frac{\#(G)}{n!}$$

if Δ is a set of $n-1$ consecutive observations. However, we know that with $n-1$ observations $\#(G_{(\Delta)}) = 1$, such that it must be the case that $\#(G) = n!$. This is only possible if $G = S_\Omega$ □

8 Fast Learning with Absence of a Common Description

We are most interested in the case where an *a priori* absence of a common description is compatible with fast learning. So far we have shown that fast learning "is the rule" and that sometimes it can indeed occur in conjunction with an *a priori* absence of a common description. One may then be led to the conjecture that there is a rich set of environments (subgroups of S_Ω) with such a cooccurrence. This section shows that this conjecture can be confirmed in a qualified sense.

We will deal with the qualification first. Our first result in this section focusses on a particular class of cooccurrences of fast learning and *a priori* absence of a common description. These are learning patterns composed of a period of incremental learning followed by a jump to a full common description. Our result on this type of learning is essentially negative; there are definite restrictions on the nature of such cooccurrences. In particular, we will show that for sufficiently large n and most $l < n$, there does not exist a group $G \leq S_\Omega$ such that Ω can be learned with l observations while $G_{(\Delta)}$ remains transitive on $\Omega \setminus \Delta$ for all Δ with $\#\Delta < l$.

In order to prove this result we need to introduce a few additional concepts from the theory of permutation groups. Call any 2-cycle a *transposition*. One can show (e.g. Rotman [1996], p.63) that for $n \geq 2$ every $g \in S_\Omega$ is a product of transpositions. If g can be factored into an even number of transpositions, then g is called an *even permutation*. The set of all even permutations in S_Ω forms a subgroup, A_Ω, that is referred to as the *alternating group* of degree n.

We also need the concept of a multiply transitive group. If G is a group acting on Ω, one can define an action on Ω^k by

$$g(\omega_1, \ldots, \omega_k) := (g(\omega_1), \ldots, g(\omega_k)).$$

Consider $\Omega^{(k)}$, that subset of Ω^k that is composed of all those k-tuples whose elements are distinct; $\Omega^{(k)}$ is G-invariant for all G and for all k. G is called *k-transitive* if G is transitive on $\Omega^{(k)}$.[8]

The following facts about k-transitive groups will be useful (e.g. Dixon and Mortimer [1996], p.33, and Wielandt [1964], p.19). For $k > 1$, k-transitivity implies $(k-1)$-transitivity. G is k-transitive on Ω if and only if G_ω is $(k-1)$-transitive on $\Omega \setminus \omega$. G is transitive if and only if it is 1-transitive. The alternating group A_Ω is $(\#\Omega - 2)$-transitive. Finally, in addition to these elementary facts about multiply transitive groups, we will make use of the following result by Wielandt [1960][9]:

Theorem 1 *(Wielandt) Let $G \leq S_\Omega$ be an 8-transitive group of finite degree. Then $G \geq A_\Omega$.*

Wielandt's proof of this result assumes what is known as the Schreier Conjecture. The Schreier conjecture in turn can be established via the classification of finite simple groups. Actually, using this classification one can strengthen the result further to show that unless a finite permutation group contains the alternating group, it is at most 5-transitive (e.g. Dixon and Mortimer [1996], p.218).

Call an observation that identifies only the observed element itself an *incremental observation*. At the other extreme are observations that lead to a simultaneous identification of all the remaining elements of Ω; those observations will be referred to as *revealing observations*. Note that when only two elements of Ω remain unidentified, then an additional observation is automatically revealing. S_Ω, for example, is learned with $n-2$ incremental observations, followed by one revealing observation. If g is an n-cycle, then $< g >$ is learned with zero incremental observations and a single revealing observation.

These two examples represent opposite ends of the spectrum of possible learning speeds. What about the intermediate ranges of the spectrum? Consider A_4, the alternating group of degree four. If we list each element of

[8] Multiply transitive groups made an early appearance in game theory in von Neumann and Morgenstern's [1947] discussion of symmetry in games.

[9] A statement, proof and discussion of this result can also be found in Dixon and Mortimer [1996], p.218

A_4 as a permutation of the column vector $(1,2,3,4)'$, then A_4 has the form

$$A_4 = \begin{array}{cccccccccccc} 1 & 2 & 3 & 4 & 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 \\ 2 & 1 & 4 & 3 & 4 & 3 & 3 & 4 & 2 & 1 & 1 & 2 \\ 3 & 4 & 1 & 2 & 2 & 4 & 1 & 3 & 4 & 2 & 3 & 1 \\ 4 & 3 & 2 & 1 & 3 & 2 & 4 & 1 & 1 & 4 & 2 & 3 \end{array}$$

If the element ω with absolute rank 2 is observed, we can first alter the standard such that this element acquires rank 1 and then focus on those permutations which fix this element. In terms of the original standard, this is the set of permutations

$$\begin{array}{ccc} 2 & 3 & 4 \\ 1 & 1 & 1 \\ 4 & 2 & 3 \\ 3 & 4 & 2 \end{array}$$

Note that with the exception of element number 2 (according to the original standard), this collection of descriptions leaves every element unidentified. The stabilizer of the element "2" is transitive on the complement of "2;" the observation "2" is incremental. The properties of multiply transitive groups imply straightforwardly that instead of "2" we could have considered any other observation and would have obtained the same result. Thus, given that A_4 expresses the players' beliefs, the first observation is incremental. Note further, that in the induced subgroup only the identity fixes any of the remaining elements. Therefore, whatever the second observation, it will be revealing.

In summary, in the case where Ω has four elements, we can find permutation groups such that all observations are either incremental or revealing and where either the first observation is revealing ($< g >$, where g is a 4-cycle), the second observations is revealing (A_4), or the third observation is revealing (S_4). Conditional on all observations being revealing or incremental, a four-element set then permits the full range of possibilities. However, as indicated before, this example is the exception when it comes to exhausting the possibilities of combining incremental and revealing observations.

If G acting on Ω can be learned with l incremental observations followed by one revealing one, we say that "Ω can be learned with l incremental observations." We saw that a four-element set could be learned with either 0, 1, or 2 incremental observations; these are all the possibilities since the $(n-1)$st observations is always revealing. According to the next proposition this state of affairs is quite exceptional.

Proposition 5 *If $\#\Omega \geq 11$, then Ω cannot be learned with l incremental observations whenever $7 \leq l \leq \#\Omega - 4$.*

Proof: In order to arrive at a contradiction, suppose that $\#\Omega \geq 11$ and that Ω can be learned with l incremental observations where l is in the specified interval. Then the group G that expresses the players' uncertainty over Ω must be $(l+1)$-transitive, i.e. still transitive on the complement of the observations after l observations, but not $(l+2)$-transitive for otherwise at least $l+1$ observations are needed to learn Ω. Thus, if $l \geq 7$, then G must be 8-transitive and therefore by Wielandt's theorem (Theorem 1), G contains the alternating group. But then G is $(\#\Omega - 2)$-transitive and therefore the pointwise stabilizer of a set Δ with no more than $\#\Omega - 3$ elements is still transitive on $\Omega \setminus \Delta$, contrary to the assumption that the $l+1$st observation is revealing (which is part of the assumption that Ω can be learned with l incremental observations). □

Note that since Theorem 1 can be strengthened, the bounds on "learning with l incremental observations" could be tightened as well.

Despite the last result, one can show that there is indeed a large set of scenarios in which an a priori absence of a common description is compatible with the possibility of fast learning. Of course, the previous result tells us that incremental observations do not play an important role in such learning. Most observations will be "partially revealing," i.e. besides identifying the observed object itself, they introduce identifying distinctions among the other objects that have not yet been observed.

Definition 1 *A group $G \leq S_\Omega$ can be learned in k steps if $G_{(\overline{\omega})} \neq e$, $\forall \overline{\omega} \in \Omega^{k-1}$ and $\exists \hat{\omega} \in \Omega^k$ such that $G_{(\hat{\omega})} = e$*

The following result shows that for any k and any lower bound on the size of Ω there is always an Ω and a subgroup of S_Ω that acts transitively on Ω and can be learned in k steps.

Proposition 6 *For any number of steps, k, and any lower bound l on $\#\Omega$, there always exists a set Ω and a group $G \leq S_\Omega$ that acts transitively on Ω and can be learned in k steps.*

Proof: Let $n = k \times l$, and for any subset K of S_Ω call the smallest subgroup of S_Ω containing K the *group generated by* K. Consider the group $H \leq S_n$ that is generated by the ("component") cycles

$$(il+1, \ldots, (i+1)l) \quad i = 0, \ldots, k-1$$

and by the cycle product

$$\prod_{j=1}^{l} (j, j+l, \ldots, j+(k-1)l).$$

To verify transitivity, first note that if $\omega = 1$ belongs to an orbit, then all elements that are moved by the component cycle $(1, \ldots, l)$ belong to the same orbit. This follows from considering compositions of cycles. Similarly, examining powers of the cycle product, it follows that if 1 belongs to an orbit, $1+l$, $1+2l, \ldots$ all belong to that orbit. Furthermore, all of these are moved by one of the different component cycles, and repeated application of those component cycles shows that for each such cycle all elements moved by that cycle belong to the orbit of $\omega = 1$. Thus all elements belong to the same orbit.

To see that k observations suffice to learn Ω, i.e. $\exists \hat{\omega}$ such that $G_{(\hat{\omega})} = e$, note that each observation ω removes those cycles from the set of remaining generators that move that element. The first observation therefore removes one component cycle *and* the cycle product. There are only $k-1$ cycles left that form the generators of G_ω. Then simply pick the remaining $k-1$ observations from different component cycles such that each of those observations eliminates one of those cycles from the set of remaining generators.

Finally, observe that $k-1$ (or fewer) observations are insufficient for learning Ω because there are k component cycles and each observation removes at most one component cycle from the set of generators. □

Intuitively, in the construction used in the proof we are stacking multiple identical cycles on top of each other, then we allow each cycle to be independently rotated and in addition we rotate the stack of cycles. This construction can be modified to obtain further classes of groups that permit k-step learning. Simply replace the component l-cycles by other permutation groups that move only the same l-elements.

9 Relation to the Literature

The problem of language learning, structure in language and creative use of language are relatively new in economics. Closest in spirit to the present paper is probably Rubinstein [1996] who is concerned with the structure of binary relations appearing in natural language. Like us, he has as one of his premises that "evolutionary forces make it more likely that the 'optimal' structures are observed [...]." He argues for certain properties that make binary relations in language more useful; among them the facility with which nameless elements in a set can be indicated, which reminds one of creative language use. He finds that this criterion of "indication-friendliness" is only satisfied by linear orderings.

For linguists of course, the questions of structure in language and how language is learned are central. Noam Chomsky's research agenda for example attempts to identify a universal language faculty, a "generative grammar." Such a grammar would account for the fact that apparently relatively few observations suffice to learn a language that is capable of generating expressions of infinitely many meanings, in particular novel expressions that have not been encountered before. Thus viewed, the generative grammar accounts for creative language use (e.g. Chomsky [1988]).

Within economics there is recent work by Segal [1996] that is related to ours. He starts with a common language but assumes that players cannot use the language to deduce a way of playing a coordination game. Instead they communicate with their common language within some organizational form. His objective is to determine the optimal organizational form (protocol, mechanism), where the likely quality of the communication outcome is traded off against a measure of communication complexity. Like in the present paper and in Rubinstein's work the focus is on characterizing efficient structures (protocols), and players are assumed to attain common knowledge of the structure via prior exposure to similar coordination problems. Segal shows that coordination by authority performs well in a wide range of circumstances and alludes to the possibility of using models of this kind to think of organizations as incomplete contracts.

Two other recent papers examine how coordination is affected or aided by structure. Chwe [1996] examines how structure affects collective action. In Chwe's model individuals share a common interest in coordinating collective action but have to rely on social networks to spread information about each individual's readiness to participate in collective action. The form of the network determines the speed at which information travels, the likelihood of collective action, and the time spent until collective action is achieved. Calvert [1991] considers Crawford and Haller's "learning to coordinate" paradigm under constantly changing conditions. He declares that "[...] the basic problem of social order involves the achievement of coordination in fundamentally *new* situations [...]," and points out the benefit to players who make use of the "common-knowledge environment" of the games.

10 Conclusion

Organizations conduct their activity in the face of a continually changing environment. It is in the very nature of novel challenges to an organization that adequate responses often have to be devised "on the spot." The members of the organization then face a chronic coordination problem and are likely to make use of any common knowledge structure that facilitates solution of novel coordination problems. This paper shows that such common knowledge structures may be of great value even if the coordination problems faced are characterized by absence of a common language. We can imagine these common knowledge structures as being established through repeated confrontations with novel but similar problems. One can perhaps think of organizations partially as reservoirs of such common knowledge structures. A successful organization is then one whose repertoire of common knowledge structures is better adapted to the changing nature of its environment.

References

AITCHISON, J. [1993], *Linguistics,* Chicago: NTC Publishing Group.

BATALI, J. [1996], "Computational Simulations of the Emergence of Grammar," University of California-San Diego, Department of Cognitive Science Working Paper .

BLUME, A. [1996], "Neighborhood Stability in Sender-Receiver Games," *Games and Economic Behavior,* **13**, 2-25.

BLUME, A., Y.-G. KIM AND J. SOBEL [1993], "Evolutionary Stability in Games of Communication," *Games and Economic Behavior,* **5**, 547-575.

CALVERT, R.L. [1991], "Elements of a Theory of Society among Rational Actors," University of Rochester, Department of Political Science Working Paper.

CHOMSKY, N. [1990], "Language and the Problem of Knowledge," in E.P. Martinich (ed.) [1996], *The Philosophy of Language,* New York: Oxford University Press.

CHWE, M. S.-Y. [1996], "Structure and Strategy in Collective Action: Communication and Coordination in Social Networks," University of Chicago, Department of Economics Working Paper.

CRAWFORD, V. AND H. HALLER [1990], "Learning how to Cooperate: Optimal Play in Repeated Coordination Games," *Econometrica,* **58**, 581-596.

DIXON, J.D. AND B. MORTIMER [1996] *Permutation Groups,* New York: Springer Verlag.

ROTMAN, J. [1996] *A First Course in Abstract Algebra,* Upper Saddle River, New Jersey: Prentice Hall.

RUBINSTEIN, A. [1991] "Comments on the Interpretation of Game Theory," *Econometrica,* **56**, 909-924.

RUBINSTEIN, A. [1996] "Why are Certain Properties of Binary Relations Relatively More Common in Natural Language?" *Econometrica,* **64**, 343-355.

SCHELLING, T.C. [1960], *The Strategy of Conflict*, Cambridge: Harvard University Press.

SEGAL, ILYA [1996] *Communication Complexity and Coordination by Authority*, University of California-Berkeley, Department of Economics Working Paper.

VON NEUMANN, J. AND O. MORGENSTERN [1947], *Theory of Games and Economic Behaviour*, Princeton: Princeton University Press.

WÄRNERYD, K. [1993], "Cheap Talk, Coordination, and Evolutionary Stability," *Games and Economic Behavior*, **5**, 532-546.

WIELANDT, H. [1960] "Über den Transitivitätsgrad von Permutationsgruppen," *Math. Z.* **74**, 297-298.

WIELANDT, H. [1964] *Finite Permutation Groups*, New York: Academic Press.

No.	Author(s)	Title
9689	T. ten Raa and E.N. Wolff	Outsourcing of Services and the Productivity Recovery in U.S. Manufacturing in the 1980s
9690	J. Suijs	A Nucleolus for Stochastic Cooperative Games
9691	C. Seidl and S.Traub	Rational Choice and the Relevance of Irrelevant Alternatives
9692	C. Seidl and S.Traub	Testing Decision Rules for Multiattribute Decision Making
9693	R.M.W.J. Beetsma and H. Jensen	Inflation Targets and Contracts with Uncertain Central Banker Preferences
9694	M. Voorneveld	Equilibria and Approximate Equilibria in Infinite Potential Games
9695	F.B.S.L.P. Janssen and A.G. de Kok	A Two-Supplier Inventory Model
9696	L. Ljungqvist and H. Uhlig	Catching up with the Keynesians
9697	A. Rustichini	Dynamic Programming Solution of Incentive Constrained Problems
9698	G.Gürkan and A.Y. Özge	Sample-Path Optimization of Buffer Allocations in a Tandem Queue - Part I: Theoretical Issues
9699	H. Huizinga	The Dual Role of Money and Optimal Financial Taxes
96100	H. Huizinga	The Taxation Implicit in Two-Tiered Exchange Rate Systems
96101	H. Norde, F. Patrone and S. Tijs	Characterizing Properties of Approximate Solutions for Optimization Problems
96102	M. Berg, A. De Waegenaere and J. Wielhouwer	Optimal Tax Reduction by Depreciation: A Stochastic Model
96103	G. van der Laan, D. Talman and Z. Yang	Existence and Approximation of Robust Stationary Points on Polytopes
96104	H. Huizinga and S.B. Nielsen	The Coordination of Capital Income and Profit Taxation with Cross-Ownership of Firms
96105	H. Degryse	The Total Cost of Trading Belgian Shares: Brussels Versus London
96106	H. Huizinga and S.B. Nielsen	The Political Economy of Capital Income and Profit Taxation in a Small Open Economy
96107	T. Dieckmann	The Evolution of Conventions with Endogenous Interactions
96108	F. de Jong and M.W.M. Donders	Intraday Lead-Lag Relationships Between the Futures-, Options and Stock Market
96109	F. Verboven	Brand Rivalry, Market Segmentation, and the Pricing of

No.	Author(s)	Title
		Optional Engine Power on Automobiles
96110	D. Granot, H. Hamers and S. Tijs	Weakly Cyclic Graphs and Delivery Games
96111	P. Aghion, P. Bolton and S. Fries	Financial Restructuring in Transition Economies
96112	A. De Waegenaere, R. Kast and A. Lapied	Non-linear Asset Valuation on Markets with Frictions
96113	R. van den Brink and P.H.M. Ruys	The Internal Organization of the Firm and its External Environment
96114	F. Palomino	Conflicting Trading Objectives and Market Efficiency
96115	E. van Damme and S. Hurkens	Endogenous Stackelberg Leadership
96116	E. Canton	Business Cycles in a Two-Sector Model of Endogenous Growth
9701	J.P.J.F. Scheepens	Collusion and Hierarchy in Banking
9702	H.G. Bloemen and E.G.F. Stancanelli	Individual Wealth, Reservation Wages and Transitions into Employment
9703	P.J.J. Herings and V.J. Vannetelbosch	Refinements of Rationalizability for Normal-Form Games
9704	F. de Jong, F.C. Drost and B.J.M. Werker	Exchange Rate Target Zones: A New Approach
9705	C. Fernández and M.F.J. Steel	On the Dangers of Modelling Through Continuous Distributions: A Bayesian Perspective
9706	M.A. Odijk, P.J. Zwaneveld, J.S. Hooghiemstra, L.G. Kroon and M. Salomon	Decision Support Systems Help Railned to Search for 'Win-Win' Solutions in Railway Network Design
9707	G. Bekaert, R.J. Hodrick and D.A. Marshall	The Implications of First-Order Risk Aversion for Asset Market Risk Premiums
9708	C. Fernández and M.F.J. Steel	Multivariate Student-*i* Regression Models: Pitfalls and Inference
9709	H. Huizinga and S.B. Nielsen	Privatization, Public Investment, and Capital Income Taxation
9710	S. Eijffinger, E. Schaling and M. Hoeberichts	Central Bank Independence: a Sensitivity Analysis
9711	H. Uhlig	Capital Income Taxation and the Sustainability of Permanent Primary Deficits

No.	Author(s)	Title
9712	M. Dufwenberg and W. Güth	Indirect Evolution Versus Strategic Delegation: A Comparison of Two Approaches to Explaining Economic Institutions
9713	H. Uhlig	Long Term Debt and the Political Support for a Monetary Union
9714	E. Charlier, B. Melenberg and A. van Soest	An Analysis of Housing Expenditure Using Semiparametric Models and Panel Data
9715	E. Charlier, B. Melenberg and A. van Soest	An Analysis of Housing Expenditure Using Semiparametric Cross-Section Models
9716	J.P. Choi and S.-S. Yi	Vertical Foreclosure with the Choice of Input Specifications
9717	J.P. Choi	Patent Litigation as an Information Transmission Mechanism
9718	H.Degryse and A. Irmen	Attribute Dependence and the Provision of Quality
9719	A. Possajennikov	An Analysis of a Simple Reinforcing Dynamics: Learning to Play an "Egalitarian" Equilibrium
9720	J. Jansen	Regulating Complementary Input Supply: Cost Correlation and Limited Liability
9721	J. ter Horst and M. Verbeek	Estimating Short-Run Persistence in Mutual Fund Performance
9722	G. Bekaert and S.F. Gray	Target Zones and Exchange Rates: An Empirical Investigation
9723	M. Slikker and A. van den Nouweland	A One-Stage Model of Link Formation and Payoff Division
9724	T. ten Raa	Club Efficiency and Lindahl Equilibrium
9725	R. Euwals, B. Melenberg and A. van Soest	Testing the Predictive Value of Subjective Labour Supply Data
9726	C. Fershtman and U. Gneezy	Strategic Delegation: An Experiment
9727	J. Potters, R. Sloof and F. van Winden	Campaign Expenditures, Contributions and Direct Endorsements: The Strategic Use of Information and Money to Influence Voter Behavior
9728	F.H. Page, Jr.	Existence of Optimal Auctions in General Environments
9729	M. Berliant and F.H. Page, Jr.	Optimal Budget Balancing Income Tax Mechanisms and the Provision of Public Goods
9730	S.C.W. Eijffinger and Willem H. Verhagen	The Advantage of Hiding Both Hands: Foreign Exchange Intervention, Ambiguity and Private Information
9731	A. Ridder, E. van der Laan and M. Salomon	How Larger Demand Variability may Lead to Lower Costs in the Newsvendor Problem
9732	K. Kultti	A Model of Random Matching and Price Formation

No.	Author(s)	Title
9733	J. Ashayeri, R. Heuts and B. Tammel	Applications of P-Median Techniques to Facilities Design Problems: an Improved Heuristic
9734	M. Dufwenberg, H. Norde, H. Reijnierse, and S. Tijs	The Consistency Principle for Set-valued Solutions and a New Direction for the Theory of Equilibrium Refinements
9735	P.P. Wakker, R.H. Thaler and A. Tversky	Probabilistic Insurance
9736	T. Offerman and J. Sonnemans	What's Causing Overreaction? An Experimental Investigation of Recency and the Hot Hand Effect
9737	R. Kabir	New Evidence on Price and Volatility Effects of Stock Option Introductions
9738	M. Das and B. Donkers	How Certain are Dutch Households about Future Income? An Empirical Analysis
9739	R.J.M. Alessie, A. Kapteyn and F. Klijn	Mandatory Pensions and Personal Savings in the Netherlands
9740	W. Güth	Ultimatum Proposals - How Do Decisions Emerge? -
9741	I. Woittiez and A. Kapteyn	Social Interactions and Habit Formation in a Model of Female Labour Supply
9742	E. Canton and H. Uhlig	Growth and the Cycle: Creative Destruction Versus Entrenchment
9743	T. Feenstra, P. Kort and A. de Zeeuw	Environmental Policy in an International Duopoly: An Analysis of Feedback Investment Strategies
9744	A. De Waegenaere and P. Wakker	Choquet Integrals with Respect to Non-Monotonic Set Functions
9745	M. Das, J. Dominitz and A. van Soest	Comparing Predicitions and Outcomes: Theory and Application to Income Changes
9746	T. Aldershof, R. Alessie and A. Kapteyn	Female Labor Supply and the Demand for Housing
9747	S.C.W. Eijffinger, M. Hoeberichts and E. Schaling	Why Money Talks and Wealth Whispers: Monetary Uncertainty and Mystique
9748	W. Güth	Boundedly Rational Decision Emergence -A General Perspective and Some Selective Illustrations-
9749	M. Lettau	Comment on 'The Spirit of Capitalism and Stock-Market Prices' by G.S. Bakshi and Z. Chen (AER, 1996)
9750	M.O. Ravn and H. Uhlig	On Adjusting the HP-Filter for the Frequency of Observations
9751	Th. v.d. Klundert and S. Smulders	Catching-Up and Regulation in a Two-Sector Small Open Economy

No.	Author(s)	Title
9752	J.P.C. Kleijnen	Experimental Design for Sensitivity Analysis, Optimization, and Validation of Simulation Models
9753	A.B.T.M. van Schaik and H.L.F. de Groot	Productivity and Unemployment in a Two-Country Model with Endogenous Growth
9754	H.L.F. de Groot and R. Nahuis	Optimal Product Variety, Scale Effects, and Growth
9755	S. Hochguertel	Precautionary Motives and Portfolio Decisions
9756	K. Kultti	Price Formation by Bargaining and Posted Prices
9757	K. Kultti	Equivalence of Auctions and Posted Prices
9758	R. Kabir	The Value Relevance of Dutch Financial Statement Numbers for Stock Market Investors
9759	R.M.W.J. Beetsma and H. Uhlig	An Analysis of the "Stability Pact"
9760	M. Lettau and H. Uhlig	Preferences, Consumption Smoothing, and Risk Premia
9761	F. Janssen and T. de Kok	The Optimal Number of Suppliers in an (s,Q) Inventory System with Order Splitting
9762	F. Janssen and T. de Kok	The Fill Rate Service Measure in an (s,Q) Inventory System with Order Splitting
9763	E. Canton	Fiscal Policy in a Stochastic Model of Endogenous Growth
9764	R. Euwals	Hours Constraints within and between Jobs
9765	A. Blume	Fast Learning in Organizations

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS